О верхней оценке математического ожидания нормы равномерно распределённого на сфере вектора и явлении концентрации равномерной меры на сфере

Édvard Gorbunov, Evgeniya Vorontsova, Aleksandr Gasnikov

Математические заметки

Том 110 выпуск 6 июнь 2021

УДК 519.21

О верхней оценке математического ожидания нормы равномерно распределённого на сфере вектора и явлении концентрации равномерной меры на сфере

Э.А. Горбунов, Е.А. Воронцова, А.В. Гасников

Рассматривается задача построения верхних оценок математического ожидания нормы равномерно распределённого на единичной евклидовой сфере вектора.

Библиография: 15 названий.

Ключевые слова: концентрация меры, равномерно распределённый на сфере вектор

1. Введение

Пусть $e \in RS_2^n(1)$ — случайный вектор, имеющий равномерное распределение на n-мерной единичной евклидовой сфере.

В настоящей работе рассматривается задача построения уточнённых верхних оценок математического ожидания нормы вектора e. От точности построения таких оценок зависят, например, оценки скорости сходимости ускоренного метода (Accelerated by Coupling Directional Search – ACDS), построенного на базе специального каплинга спусков по направлению в форме градиентного спуска и метода зеркального спуска [1].

Первые такие оценки были получены в 2014 году в [2], статья опубликована в 2016 году в [3]. Кроме того, независимо от работы [3] в 2015 г. похожие оценки были сделаны в [4], опубликовано в 2017 г. в [5]; и, также независимо, в 2015 г. одна из возможных оценок математического ожидания нормы равномерно распределённого на единичной евклидовой сфере вектора была получена в [6].

Основным результатом данной работы является теорема 1.

2. Постановка задачи и формулировка результата

Пусть задан некоторый (неслучайный) вектор s с единичной евклидовой сферы. Не умаляя общности, мы будем считать, что вектор s направлен вдоль первой координатной оси (если это не так, то мы можем перейти к нужному базису). Тогда

Работа А.В. Гасникова и Э.А. Горбунова поддержана грантом РФФИ 18-31-20005 мол_а_вед. Работа Е.А. Воронцовой поддержана грантом РФФИ 18-29-03071.

с вероятностью хотя бы $1-\frac{2}{c}e^{-\frac{c^2}{2}}$ будет выполнено неравенство $|\langle s, e\rangle| \leqslant \frac{c}{\sqrt{n-1}}$ (см. теорему 2.7 и рисунок 2.2 из [7] и [8]). То есть, если взять $c=10$, то получим, что с большой вероятностью выполнено неравенство $\langle s, e\rangle^2 \leqslant \frac{100}{n}$ (множество, на котором $\langle s, e\rangle^2 \leqslant \frac{100}{n}$, обозначим через A_s; как мы видим, при достаточно больших n вероятностная мера множества A_s велика). Кроме того, можно показать, что $\mathbb{E}[\langle s, e\rangle^2] = \frac{1}{n}$ (см., например, лемму B.10 из [9]).

Рассмотрим ∞-норму, которая для произвольного вектора $x \in \mathbb{R}^n$ задаётся формулой $\|x\|_\infty = \max\limits_{1\leqslant i\leqslant n} |x_i|$, где $x = (x_1, x_2, \ldots, x_n)^\top$. Заметим, что функция $f(e) = \|e\|_\infty$ является липшицевой с константой 1 в евклидовой норме. Рассмотрим константу M_f такую, что $\mathbb{P}_e\left\{f(e) \geqslant M_f\right\} \geqslant \frac{1}{2}$ и $\mathbb{P}_e\left\{f(e) \leqslant M_f\right\} \geqslant \frac{1}{2}$. Тогда верно неравенство (см. [10], [11])

$$\mathbb{P}_e\left\{|f(e) - M_f| > t\right\} \leqslant 4e^{-\frac{t^2}{4}}, \quad t > 0.$$

Это означает, что случайная величина $\|e\|_\infty$ принимает очень близкие к $\mathbb{E}[\|e\|_\infty]$ (M_f и $\mathbb{E}[f(e)]$ асимптотически близки, см. [12]) значения на множестве достаточно большой меры. Кроме того, можно показать, что максимальная по модулю компонента вектора e с вероятностью не меньше $1 - \frac{1}{n\sqrt{n}}$ принимает значения по модулю меньшие $\frac{2\sqrt{\ln n}}{\sqrt{n-1}}$ (множество, на котором $\|e\|_\infty \leqslant \frac{2\sqrt{\ln n}}{\sqrt{n-1}}$, обозначим через B_∞). Тогда $\mathbb{E}[\langle s, e\rangle^2\|e\|_\infty^2]$ близко к среднему значению случайной величины $\langle s, e\rangle^2\|e\|_\infty^2$ на множестве $A_e \cap B_\infty$ (чья вероятностная мера по-прежнему велика), на котором она не превосходит $400\ln n/n^2$. Константа в этой оценке сильно завышена и она уточняется далее в Теореме 1 (причём не только для ∞-нормы). Однако такого рода рассуждения, вытекающие из явления концентрации равномерной меры на сфере, поясняют причины возникновения такой оценки, а также её целесообразность в терминах вхождения размерности пространства n.

Сформулируем и докажем достаточно известный факт, заключающийся в том, что векторная q-норма является невозрастающей функцией от q для любого фиксированного вектора.

Лемма 1. *Для любого вектора $x \in \mathbb{R}^n$ (и для любого $n \in \mathbb{N}$) выполнено неравенство:*

$$\|x\|_{p_1} \leqslant \|x\|_{p_2}, \tag{2.1}$$

где $p_1 \geqslant p_2$ и под знаком $\|\cdot\|_q$ понимается векторная q-норма (норма Гёльдера с показателем q).

Доказательство. *Не умаляя общности, будем считать, что все компоненты вектора x являются ненулевыми (если вектор x ненулевой, то его норма равна норме его подвектора меньшей размерности, полученного удалением нулевых компонент, и, соответственно, можно рассматривать этот подвектор; если же вектор $x = 0$, то неравенство (2.1) тоже верно). Пусть $g_x(p) \stackrel{def}{=} \ln\|x\|_p =$*

$\ln\left(\sum\limits_{k=1}^{n}|x_k|^p\right)^{\frac{1}{p}} = \frac{1}{p}\ln\left(\sum\limits_{k=1}^{n}|x_k|^p\right)$. *Тогда*

$$\frac{dg_x(p)}{dp} = -\frac{1}{p^2}\ln\left(\sum_{k=1}^{n}|x_k|^p\right) + \frac{1}{p}\cdot\frac{\sum\limits_{k=1}^{n}\ln(|x_k|)\cdot|x_k|^p}{\sum\limits_{k=1}^{n}|x_k|^p}.$$

Так как $\ln y$ *— вогнутая по* y *функция, то по неравенству Йенсена получаем*

$$\begin{aligned}\frac{dg_x(p)}{dp} &\leqslant \frac{1}{p}\ln\left(\sum_{k=1}^{n}|x_k|^p\right)^{-\frac{1}{p}} + \frac{1}{p}\ln\left(\sum_{k=1}^{n}|x_k|\cdot\frac{|x_k|^p}{\sum\limits_{k=1}^{n}|x_k|^p}\right)\\ &= \frac{1}{p}\ln\left(\frac{\sum\limits_{k=1}^{n}|x_k|^{p+1}}{\left(\sum\limits_{k=1}^{n}|x_k|^p\right)^{\frac{p+1}{p}}}\right) \leqslant \frac{1}{p}\ln\left(\frac{\sum\limits_{k=1}^{n}|x_k|^{p+1}}{\sum\limits_{k=1}^{n}(|x_k|^p)^{\frac{p+1}{p}}}\right)\\ &= 0,\end{aligned}$$

то есть функция $g_x(p)$ *— невозрастающая функция на* $[1,+\infty)$. *Лемма доказана.*

Имеет место следующая теорема, являющаяся следствием явления концентрации равномерной меры на сфере вокруг экватора (см. также [13]; северный полюс задается градиентом $\nabla f(x)$).

Теорема 1. *Пусть* $e \in RS_2^n(1)$, $n \geqslant 8$, $s \in \mathbb{R}^n$, *тогда*

$$\mathbb{E}[||e||_q^2] \leqslant \min\{q-1,\, 16\ln n - 8\}n^{\frac{2}{q}-1}, \quad 2 \leqslant q \leqslant \infty \tag{2.2}$$

$$\mathbb{E}[\langle s,\, e\rangle^2||e||_q^2] \leqslant \sqrt{3}||s||_2^2\min\{2q-1, 32\ln n - 8\}n^{\frac{2}{q}-2}, \quad 2 \leqslant q \leqslant \infty, \tag{2.3}$$

где под знаком $||\cdot||_q$ *понимается векторная* q*-норма (норма Гёльдера с показателем* q*).*

Доказательство. *Докажем вспомогательное неравенство:*

$$\mathbb{E}[||e||_q^2] \leqslant (q-1)n^{\frac{2}{q}-1}, \quad 2 \leqslant q < \infty. \tag{2.4}$$

Во-первых,

$$\mathbb{E}[||e||_q^2] = \mathbb{E}\left[\left(\sum_{k=1}^{n}|e_k|^q\right)^{\frac{2}{q}}\right] \overset{①}{\leqslant} \left(\mathbb{E}\left[\sum_{k=1}^{n}|e_k|^q\right]\right)^{\frac{2}{q}} \overset{②}{=} \left(n\mathbb{E}[|e_2|^q]\right)^{\frac{2}{q}}, \tag{2.5}$$

где ① выполнено в силу вероятностного неравенства Йенсена (функция $\varphi(x) = x^{\frac{2}{q}}$ *является вогнутой, так как* $q \geqslant 2$*), а переход ② корректен в силу линейности математического ожидания и одинаковой распределённости компонент вектора* e*.*

Во-вторых, воспользуемся тем фактом (лемма Пуанкаре, см., например, [14; п. 6.3]*), что*

$$e \overset{d}{=} \frac{\xi}{\sqrt{\xi_1^2 + \cdots + \xi_n^2}}, \tag{2.6}$$

где $\xi = (\xi_1, \xi_2, \ldots, \xi_n)^\top$ *— n-мерный гауссовский случайный вектор с нулевым математическим ожиданием и единичной ковариационной матрицей, а* $\stackrel{d}{=}$ *обозначает равенство по распределению. Тогда*

$$\begin{aligned}\mathbb{E}[|e_2|^q] &= \mathbb{E}\left[\tfrac{|\xi_2|^q}{\left(\xi_1^2+\ldots+\xi_n^2\right)^{\frac{q}{2}}}\right] \\ &= \int\limits_{\mathbb{R}^n}\cdots\int |x_2|^q\left(\sum_{k=1}^n x_k^2\right)^{-\frac{q}{2}}\cdot\tfrac{1}{(2\pi)^{\frac{n}{2}}}\cdot\exp\left(-\tfrac{1}{2}\sum_{k=1}^n x_k^2\right)dx_1\ldots dx_n.\end{aligned}$$

Перейдём к сферическим координатам:

$$\begin{aligned}x_1 &= r\cos\varphi\sin\theta_1\ldots\sin\theta_{n-2},\\ x_2 &= r\sin\varphi\sin\theta_1\ldots\sin\theta_{n-2},\\ x_3 &= r\cos\theta_1\sin\theta_2\ldots\sin\theta_{n-2},\\ x_4 &= r\cos\theta_2\sin\theta_3\ldots\sin\theta_{n-2},\\ \ldots & \\ x_n &= r\cos\theta_{n-2},\\ & r>0,\ \varphi\in[0,2\pi),\ \theta_i\in[0,\pi],\ i=\overline{1,n-2},\end{aligned}$$

якобиан преобразования координат равен

$$\det\left(\frac{\partial(x_1,\ldots,x_n)}{\partial(r,\varphi,\theta_1,\theta_2,\ldots,\theta_{n-2})}\right) = r^{n-1}\sin\theta_1(\sin\theta_2)^2\ldots(\sin\theta_{n-2})^{n-2}.$$

Тогда математическое ожидание $\mathbb{E}[|e_2|^q]$ *можно записать в виде:*

$$\begin{aligned}\mathbb{E}[|e_2|^q] &= \int\limits_{\substack{r>0,\,\varphi\in[0,\,2\pi),\\ \theta_i\in[0,\pi],\,i=\overline{1,n-2}}}\cdots\int r^{n-1}|\sin\varphi|^q|\sin\theta_1|^{q+1}|\sin\theta_2|^{q+2}\ldots|\sin\theta_{n-2}|^{q+n-2}\\ &\cdot\frac{e^{-\frac{r^2}{2}}}{(2\pi)^{\frac{n}{2}}}dr\ldots d\theta_{n-2}\\ &= \frac{1}{(2\pi)^{\frac{n}{2}}}I_r\cdot I_\varphi\cdot I_{\theta_1}\cdot I_{\theta_2}\cdot\ldots\cdot I_{\theta_{n-2}},\end{aligned}$$

где

$$\begin{aligned}I_r &= \int\limits_0^{+\infty} r^{n-1}e^{-\frac{r^2}{2}}dr,\\ I_\varphi &= \int\limits_0^{2\pi}|\sin\varphi|^q d\varphi = 2\int\limits_0^{\pi}|\sin\varphi|^q d\varphi,\\ I_{\theta_i} &= \int\limits_0^{\pi}|\sin\theta_i|^{q+i}d\theta_i,\ i=\overline{1,\,n-2}.\end{aligned}$$

Вычислим эти интегралы. Начнём с I_r*:*

$$I_r = \int\limits_0^{+\infty} r^{n-1}e^{-\frac{r^2}{2}}dr = /\text{замена}\, r=\sqrt{2t}/ = \int\limits_0^{+\infty}(2t)^{\frac{n}{2}-1}e^{-t}dt = 2^{\frac{n}{2}-1}\Gamma(\tfrac{n}{2}).$$

Чтобы вычислить остальные интегралы, будет полезно рассмотреть следующий интеграл ($\alpha>0$)*:*

$$\begin{aligned}\int\limits_0^{\pi}|\sin\varphi|^\alpha d\varphi &= 2\int\limits_0^{\frac{\pi}{2}}|\sin\varphi|^\alpha d\varphi = 2\int\limits_0^{\frac{\pi}{2}}(\sin^2\varphi)^{\frac{\alpha}{2}}d\varphi = /\text{замена}\, t=\sin^2\varphi/\\ &= \int\limits_0^1 t^{\frac{\alpha-1}{2}}(1-t)^{-\frac{1}{2}}dt = B(\tfrac{\alpha+1}{2},\,\tfrac{1}{2}) = \frac{\Gamma(\frac{\alpha+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{\alpha+2}{2})} = \sqrt{\pi}\frac{\Gamma(\frac{\alpha+1}{2})}{\Gamma(\frac{\alpha+2}{2})}.\end{aligned}$$

Отсюда получаем, что

$$\begin{aligned}
\mathbb{E}[|e_2|^q] &= \tfrac{1}{(2\pi)^{\frac{n}{2}}} I_r \cdot I_\varphi \cdot I_{\theta_1} \cdot I_{\theta_2} \cdot \ldots \cdot I_{\theta_{n-2}} \\
&= \tfrac{2^{\frac{n}{2}-1}}{(2\pi)^{\frac{n}{2}}} \cdot \Gamma(\tfrac{n}{2}) \cdot 2 \tfrac{\sqrt{\pi}\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+2}{2})} \cdot \tfrac{\sqrt{\pi}\Gamma(\frac{q+2}{2})}{\Gamma(\frac{q+3}{2})} \cdot \tfrac{\sqrt{\pi}\Gamma(\frac{q+3}{2})}{\Gamma(\frac{q+4}{2})} \cdot \ldots \cdot \tfrac{\sqrt{\pi}\Gamma(\frac{q+n-1}{2})}{\Gamma(\frac{q+n}{2})} \\
&= \tfrac{1}{\sqrt{\pi}} \cdot \tfrac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})}.
\end{aligned} \tag{2.7}$$

Покажем, что $\forall\, q \geqslant 2$

$$\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})} \leqslant \left(\frac{q-1}{n}\right)^{\frac{q}{2}}. \tag{2.8}$$

Сначала убедимся, что неравенство (2.8) *выполнено для* $q = 2$ *(и произвольного* n*):*

$$\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{2+1}{2})}{\Gamma(\frac{2+n}{2})} - \frac{1}{n} = \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2}) \cdot \frac{1}{2}\Gamma(\frac{1}{2})}{\frac{n}{2}\Gamma(\frac{n}{2})} - \frac{1}{n} = \frac{1}{n} - \frac{1}{n} = 0 \leqslant 0.$$

Рассмотрим функцию (вообще говоря, двух аргументов)

$$f_n(q) = \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})} - \left(\frac{q-1}{n}\right)^{\frac{q}{2}}$$

при $q \geqslant 2$. *Также введём в рассмотрение функцию* $\psi(x) = \frac{d(\ln(\Gamma(x)))}{dx}$ *при* $x > 0$ *(дигамма-функция). Для гамма-функции выполняется тождество*

$$\Gamma(x+1) = x\Gamma(x),\ x > 0.$$

Возьмём от этого тождества логарифм и продифференцируем по x*:*

$$\begin{gathered}
\ln\Gamma(x+1) = \ln\Gamma(x) + \ln x, \\
\tfrac{d(\ln(\Gamma(x+1)))}{dx} = \tfrac{d(\ln(\Gamma(x)))}{dx} + \tfrac{1}{x},
\end{gathered}$$

что можно записать через дигамма-функцию:

$$\psi(x+1) = \psi(x) + \frac{1}{x}. \tag{2.9}$$

Покажем, что дигамма-функция возрастает при $x > 0$. *Для этого докажем неравенство:*

$$\left(\Gamma'(x)\right)^2 < \Gamma(x)\Gamma''(x). \tag{2.10}$$

Действительно,

$$\begin{aligned}
\left(\Gamma'(x)\right)^2 &= \left(\int\limits_0^{+\infty} e^{-t} \ln t \cdot t^{x-1} dt\right)^2 \\
&\overset{①}{<} \int\limits_0^{+\infty} \left(e^{-\frac{t}{2}} t^{\frac{x-1}{2}}\right)^2 dt \cdot \int\limits_0^{+\infty} \left(e^{-\frac{t}{2}} t^{\frac{x-1}{2}} \ln t\right)^2 dt \\
&= \underbrace{\int\limits_0^{+\infty} e^{-t} t^{x-1} dt}_{\Gamma(x)} \cdot \underbrace{\int\limits_0^{+\infty} e^{t} t^{x-1} \ln^2 t dt}_{\Gamma''(x)},
\end{aligned}$$

где ① следует из неравенства Коши-Буняковского (причём неравенство строгое, ибо функции $e^{-\frac{t}{2}}t^{\frac{x-1}{2}}$ *и* $e^{-\frac{t}{2}}t^{\frac{x-1}{2}}\ln t$ *линейно независимы). Из неравенства* (2.10) *следует, что*

$$\frac{d^2(\ln\Gamma(x))}{dx^2} = \left(\frac{\Gamma'(x)}{\Gamma(x)}\right)' = \frac{\Gamma''(x)}{\Gamma(x)} - \frac{\left(\Gamma'(x)\right)^2}{\left(\Gamma(x)\right)^2} \overset{(2.10)}{>} 0,$$

то есть дигамма-функция возрастает.

Теперь покажем, что $f_n(q)$ *убывает на отрезке* $[2,+\infty)$*. Для этого достаточно рассмотреть* $\ln(f(q))$*:*

$$\begin{aligned}\ln(f_n(q)) &= \ln\left(\frac{\Gamma(\frac{n}{2})}{\sqrt{\pi}}\right) + \ln\left(\Gamma\left(\tfrac{q+1}{2}\right)\right) - \ln\left(\Gamma\left(\tfrac{q+n}{2}\right)\right) - \tfrac{q}{2}\left(\ln(q-1) - \ln n\right),\\ \tfrac{d(\ln(f_n(q)))}{dq} &= \tfrac{1}{2}\psi\left(\tfrac{q+1}{2}\right) - \tfrac{1}{2}\psi\left(\tfrac{q+n}{2}\right) - \tfrac{1}{2}\ln(q-1) - \tfrac{q}{2(q-1)} + \tfrac{1}{2}\ln n.\end{aligned}$$

Покажем, что $\frac{d(\ln(f_n(q)))}{dq} < 0$ *при* $q \geqslant 2$*. Пусть* $k = \lfloor\frac{n}{2}\rfloor$ *(ближайшее целое число, не превосходящее* $\frac{n}{2}$*). Тогда* $\psi\left(\frac{q+n}{2}\right) > \psi\left(k-1+\frac{q+1}{2}\right)$ *и* $\ln n \leqslant \ln(2k+1)$*, откуда следует, что*

$$\begin{aligned}\tfrac{d(\ln(f_n(q)))}{dq} &< \tfrac{1}{2}\left(\psi\left(\tfrac{q+1}{2}\right) - \psi\left(k-1+\tfrac{q+1}{2}\right)\right) - \tfrac{1}{2}\ln(q-1) - \tfrac{q}{2(q-1)} + \tfrac{1}{2}\ln(2k+1)\\ &\overset{(2.9)}{=} \tfrac{1}{2}\left(\psi\left(\tfrac{q+1}{2}\right) - \sum_{i=1}^{k-1}\frac{1}{\frac{q+1}{2}+k-i-1} - \psi\left(\tfrac{q+1}{2}\right)\right) - \tfrac{q}{2(q-1)} + \tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right)\\ &\overset{①}{\leqslant} -\tfrac{1}{2}\sum_{i=1}^{k-1}\tfrac{2}{q-1+2k-2i} - \tfrac{1}{q-1} + \tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right)\\ &= -\tfrac{1}{2}\left(\tfrac{2}{q-1} + \tfrac{2}{q+1} + \tfrac{2}{q+3} + \ldots + \tfrac{2}{q+2k-3}\right) + \tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right)\\ &\overset{②}{<} -\tfrac{1}{2}\ln\left(\tfrac{q+2k-1}{q-1}\right) + \tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right) \overset{③}{\leqslant} -\tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right) + \tfrac{1}{2}\ln\left(\tfrac{2k+1}{q-1}\right) = 0,\end{aligned}$$

где ① и ③ выполнены в силу неравенства $q \geqslant 2$*, ② следует из оценки сверху интеграла от функции* $\frac{1}{x}$ *интегралом от её верхней ступенчатой мажоранты* $g(x) = \frac{1}{q-1+2i}$*,* $x \in [q-1+2i, q-1+2i+2]$*,* $i = \overline{0, 2k-1}$*:*

$$\frac{2}{q-1} + \frac{2}{q+1} + \frac{2}{q+3} + \ldots + \frac{2}{q+2k-3} > \int\limits_{q-1}^{q+2k-1}\frac{1}{x}dx = \ln\left(\frac{q+2k-1}{q-1}\right).$$

Итак, мы показали, что $\frac{d(\ln(f_n(q)))}{dq} < 0$ *для* $q \geqslant 2$ *и произвольного натурального* n*. Следовательно, для любого фиксированного* n *функция* $f_n(q)$ *убывает по* q*, а значит,* $f_n(q) \leqslant f_n(2) = 0$*, то есть справедливо неравенство* (2.8)*. Отсюда и из* (2.5)*,* (2.7) *получаем, что для любого* $q \geqslant 2$

$$\mathbb{E}[||e||_q^2] \overset{(2.5)}{\leqslant} \left(n\mathbb{E}[|e_2|^q]\right)^{\frac{2}{q}} \overset{(2.7),(2.8)}{\leqslant} (q-1)n^{\frac{2}{q}-1}. \tag{2.11}$$

Неравенство (2.11) *нет смысла использовать при больших* q *(относительно* n*). Рассмотрим правую часть неравенства* (2.11) *как функцию* q *и найдём её минимум при* $q \geqslant 2$*. Рассмотрим* $h_n(q) = \ln(q-1) + \left(\frac{2}{q}-1\right)\ln n$ *(логарифм правой части* (2.11)*). Производная* $h(q)$*:*

$$\begin{aligned}\tfrac{dh(q)}{dq} &= \tfrac{1}{q-1} - \tfrac{2\ln n}{q^2},\\ \tfrac{1}{q-1} - \tfrac{2\ln n}{q^2} &= 0,\\ q^2 - 2q\ln n + 2\ln n &= 0.\end{aligned}$$

Если $n \geqslant 8$, то точка минимума на множестве $[2, +\infty)$ есть

$$q_0 = \ln n \left(1 + \sqrt{1 - \frac{2}{\ln n}}\right)$$

(в случае $n \leqslant 7$ оказывается, что $q_0 = 2$; везде далее считаем, что $n \geqslant 8$). Поэтому для всех $q > q_0$ более точная оценка будет следующей:

$$\begin{aligned} \mathbb{E}[||e||_q^2] &\overset{①}{<} \mathbb{E}[||e||_{q_0}^2] \overset{(2.11)}{\leqslant} (q_0 - 1)n^{\frac{2}{q_0}-1} \overset{②}{\leqslant} (2\ln n - 1)n^{\frac{2}{\ln n}-1} \\ &= (2\ln n - 1)e^2\tfrac{1}{n} \leqslant (16\ln n - 8)\tfrac{1}{n} \leqslant (16\ln n - 8)n^{\frac{2}{q}-1}, \end{aligned} \tag{2.12}$$

где ① верно в силу Леммы 1, ② следует из неравенств $q_0 \leqslant 2\ln n$, $q_0 \geqslant \ln n$. Объединяя оценки (2.11) и (2.12), получаем неравенство (2.2).

Теперь перейдём к доказательству неравенства (2.3). Во-первых, получим оценку для $\sqrt{\mathbb{E}[||e||_q^4]}$. В силу вероятностного неравенства Йенсена ($q \geqslant 2$)

$$\begin{aligned} \mathbb{E}[||e||_q^4] &= \mathbb{E}\left[\left(\left(\sum_{k=1}^n |e_k|^q\right)^2\right)^{\frac{2}{q}}\right] \leqslant \left(\mathbb{E}\left[\left(\sum_{k=1}^n |e_k|^q\right)^2\right]\right)^{\frac{2}{q}} \\ &\overset{①}{\leqslant} \left(\mathbb{E}\left[\left(n\sum_{k=1}^n |e_k|^{2q}\right)\right]\right)^{\frac{2}{q}} \overset{②}{=} \left(n^2\mathbb{E}[|e_2|^{2q}]\right)^{\frac{2}{q}} \\ &\overset{(2.7),(2.8)}{\leqslant} n^{\frac{4}{q}}\left(\left(\tfrac{2q-1}{n}\right)^{\frac{2q}{2}}\right)^{\frac{2}{q}} = (2q-1)^2 n^{\frac{4}{q}-2}, \end{aligned}$$

где ① следует из неравенства $\left(\sum\limits_{k=1}^n x_k\right)^2 \leqslant n\sum\limits_{k=1}^n x_k^2$ для $x_1, x_2, \ldots, x_n \in \mathbb{R}$, а ② есть следствие линейности математического ожидания и одинаковой распределённости компонент вектора e. Отсюда получаем оценку

$$\sqrt{\mathbb{E}[||e||_q^4]} \leqslant (2q-1)n^{\frac{2}{q}-1}. \tag{2.13}$$

Рассмотрим правую часть неравенства (2.13) как функцию q и найдём её минимум при $q \geqslant 2$. Рассмотрим $h_n(q) = \ln(2q-1) + \left(\frac{2}{q} - 1\right)\ln n$ (логарифм правой части (2.13)). Производная $h(q)$:

$$\begin{aligned} \frac{dh(q)}{dq} &= \frac{2}{2q-1} - \frac{2\ln n}{q^2}, \\ \frac{2}{2q-1} - \frac{2\ln n}{q^2} &= 0, \\ q^2 - 2q\ln n + \ln n &= 0. \end{aligned}$$

Если $n \geqslant 3$, то точка минимума на множестве $[2, +\infty)$ есть

$$q_0 = \ln n \left(1 + \sqrt{1 - \frac{1}{\ln n}}\right)$$

(в случае $n \leqslant 2$ оказывается, что $q_0 = 2$; везде далее считаем, что $n \geqslant 3$). Поэтому для всех $q > q_0$ более точная оценка будет следующей:

$$\begin{aligned} \sqrt{\mathbb{E}[||e||_q^4]} &\overset{①}{<} \sqrt{\mathbb{E}[||e||_{q_0}^4]} \overset{(2.13)}{\leqslant} (2q_0 - 1)n^{\frac{2}{q_0}-1} \overset{②}{\leqslant} (4\ln n - 1)n^{\frac{2}{\ln n}-1} \\ &= (4\ln n - 1)e^2\tfrac{1}{n} \leqslant (32\ln n - 8)\tfrac{1}{n} \leqslant (32\ln n - 8)n^{\frac{2}{q}-1}, \end{aligned} \tag{2.14}$$

где ① верно в силу неравенства $||e||_q < ||e||_{q_0}$ *для* $q > q_0$, *② следует из неравенств* $q_0 \leqslant 2\ln n$, $q_0 \geqslant \ln n$. *Объединяя оценки* (2.13) *и* (2.14), *получаем неравенство*

$$\sqrt{\mathbb{E}[||e||_q^4]} \leqslant \min\{2q-1, 32\ln n - 8\} n^{\frac{2}{q}-1}. \tag{2.15}$$

Теперь найдём $\mathbb{E}[\langle s, e\rangle^4]$, *где* $s \in \mathbb{R}^n$ *— некоторый вектор. Пусть* $S_n(r)$ *— площадь поверхности* n*-мерной евклидовой сферы радиуса* n, $d\sigma(e)$ *— ненормированная равномерная мера на* n*-мерной евклидовой сфере. В данных обозначениях* $S_n(r) = S_n(1)r^{n-1}$, $\frac{S_{n-1}(1)}{S_n(1)} = \frac{n-1}{n\sqrt{\pi}}\frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})}$. *Кроме того, пусть* φ *— угол между* s *и* e. *Тогда*

$$\begin{aligned} \mathbb{E}[\langle s, e\rangle^4] &= \tfrac{1}{S_n(1)} \int\limits_S \langle s, e\rangle^4 d\sigma(\varphi) = \tfrac{1}{S_n(1)} \int\limits_0^\pi ||s||_2^4 \cos^3\varphi S_{n-1}(\sin\varphi) d\varphi \\ &= ||s||_2^4 \tfrac{S_{n-1}(1)}{S_n(1)} \int\limits_0^\pi \cos^4\varphi \sin^{n-2}\varphi d\varphi \\ &= ||s||_2^4 \cdot \tfrac{n-1}{n\sqrt{\pi}} \tfrac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})} \int\limits_0^\pi \cos^4\varphi \sin^{n-2}\varphi d\varphi. \end{aligned} \tag{2.16}$$

Отдельно вычислим интеграл:

$$\begin{aligned} \int\limits_0^\pi \cos^4\varphi \sin^{n-2}\varphi d\varphi &= 2\int\limits_0^{\frac{\pi}{2}} \cos^4\varphi \sin^{n-2}\varphi d\varphi = /\textit{замена}\, t = \sin^2\varphi/ \\ &= \int\limits_0^{\frac{\pi}{2}} t^{\frac{n-3}{2}} (1-t)^{\frac{3}{2}} dt = B(\tfrac{n-1}{2}, \tfrac{5}{2}) = \tfrac{\Gamma(\frac{5}{2})\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n+4}{2})} \\ &= \tfrac{\frac{3}{2}\cdot\frac{1}{2}\Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})}{\frac{n+2}{2}\cdot\Gamma(\frac{n+2}{2})} = \tfrac{3}{n+2} \cdot \tfrac{\sqrt{\pi}\Gamma(\frac{n-1}{2})}{2\Gamma(\frac{n+2}{2})}. \end{aligned}$$

Отсюда и из (2.16) *получаем, что*

$$\begin{aligned} \mathbb{E}[\langle s, e\rangle^4] &= ||s||_2^4 \cdot \tfrac{n-1}{n\sqrt{\pi}} \tfrac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})} \cdot \tfrac{3}{n+2} \cdot \tfrac{\sqrt{\pi}\Gamma(\frac{n-1}{2})}{2\Gamma(\frac{n+2}{2})} \\ &= ||s||_2^4 \cdot \tfrac{3(n-1)}{2n(n+2)} \cdot \tfrac{\Gamma(\frac{n-1}{2})}{\frac{n-1}{2}\Gamma(\frac{n-1}{2})} = \tfrac{3||s||_2^4}{n(n+2)} \overset{①}{\leqslant} \tfrac{3||s||_2^4}{n^2}. \end{aligned} \tag{2.17}$$

Чтобы доказать неравенство (2.3), *осталось воспользоваться* (2.15), (2.17) *и неравенством Коши-Буняковского* ($(\mathbb{E}[XY])^2 \leqslant \mathbb{E}[X^2]\cdot\mathbb{E}[Y^2]$)*:*

$$\mathbb{E}[\langle s, e\rangle^2 ||e||_q^2] \overset{①}{\leqslant} \sqrt{\mathbb{E}[\langle s, e\rangle^4]\cdot\mathbb{E}[||e||_q^4]} \leqslant \sqrt{3}||s||_2^2 \min\{2q-1, 32\ln n - 8\} n^{\frac{2}{q}-2}.$$

Теорема доказана.

3. Вычислительные эксперименты

Для уточнения констант в верхних оценках теоремы 1 были проведены вычислительные эксперименты. При генерации случайных векторов, равномерно распределённых на поверхности n-мерной евклидовой сферы, использовалась лемма Пуанкаре (см. (2.6)) о том, что компоненты любого вектора e, имеющего такое распределение, можно представлять как отношения $\frac{e_k}{\sqrt{e_1^2+\ldots+e_n^2}}$, где все e_k, $k = 1, 2, \ldots$ —

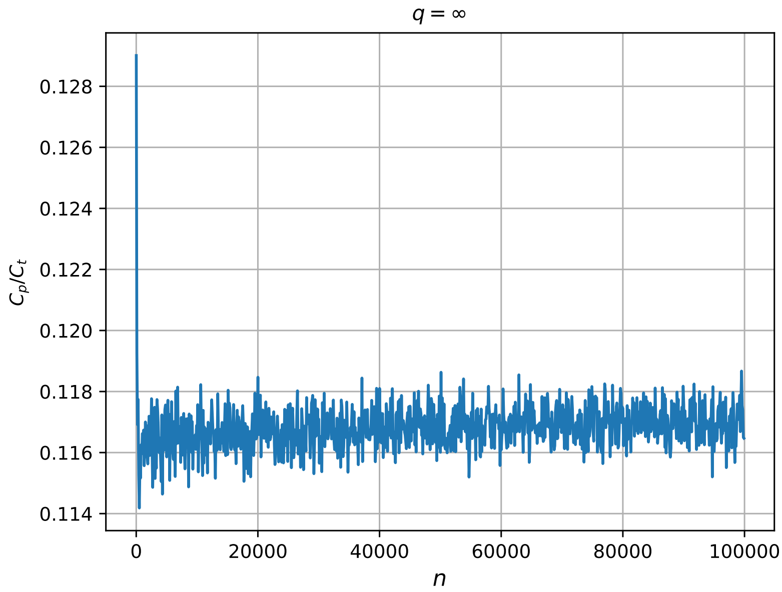

Рис. 1. Уточнение константы в оценке (2.2), n — размерность пространства, $C_t = 16\ln n - 8$

независимые одинаково распределённые случайные величины, имеющие стандартное нормальное распределение $N(0,\,1)$.

На рис. 1 приведены результаты эксперимента по оценке математического ожидания ∞-нормы векторов $e \in RS_2^n(1)$. По теореме 1 при $q = \infty$ неравенство (2.2) имеет вид

$$\mathbb{E}[||e||_\infty^2] \leqslant C_t n^{-1},$$

где $C_t = 16\ln n - 8$. Эти же константы (назовём их в этом случае C_p) можно оценить практически, путём вычисления $\mathbb{E}[||e||_\infty^2]$ методом Монте-Карло. Это было сделано для n от 10 до 10^5, и на рис. 1 приведено отношение C_p/C_t для разных n. Оказалось, что отношение с ростом n не меняется, что значит, что теоретическая оценка с точностью до константы верна.

Такие же эксперименты были проведены и для оценки константы

$$C_t = \sqrt{3}||s||_2^2 \min\{2q - 1,\, 32\ln n - 8\}$$

в неравенстве (2.3) при $q = \infty$:

$$\mathbb{E}[\langle s,\, e\rangle^2 ||e||_q^2] \leqslant C_t n^{-2}.$$

Результаты экспериментов (см. рис. 2) также подтверждают, что теоретическая оценка C_t с точностью до константы верна.

Код на языке Python всех вычислительных экспериментов выложен на Github [15].

4. Благодарности

Авторы выражают благодарность Павлу Евгеньевичу Двуреченскому за помощь в работе.

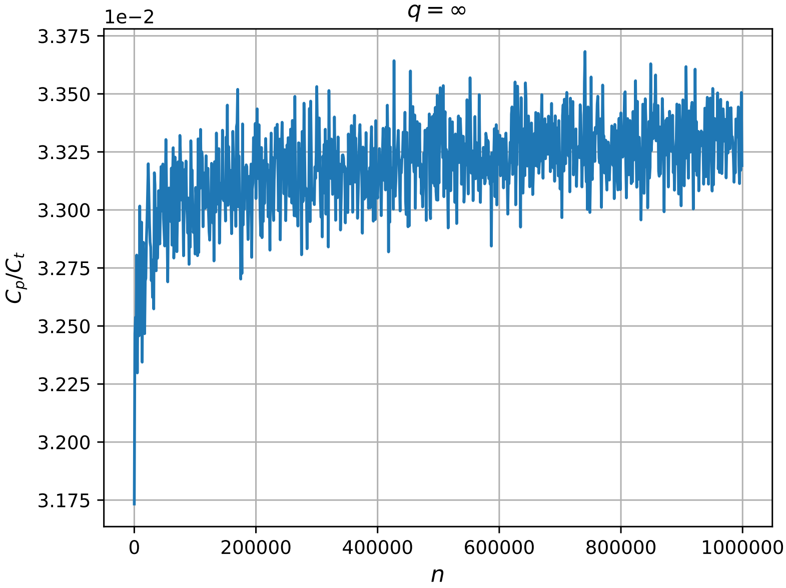

Рис. 2. Уточнение константы в оценке (2.3), n — размерность пространства, $C_t = \sqrt{3}||s||_2^2 \min\{2q-1,\, 32\ln n - 8\}$

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

[1] Е. А. Воронцова, А. В. Гасников, Э. А. Горбунов, *Ускоренные спуски по случайному направлению с неевклидовой прокс-структурой*, arXiv preprint arXiv:1710.00162.

[2] A. Gasnikov, A. Lagunovskaya, I. Usmanova, F. Fedorenko, *Gradient-free prox-methods with inexact oracle for stochastic convex optimization problems on a simplex*, arXiv preprint arXiv:1412.3890.

[3] А. В. Гасников, А. А. Лагуновская, И. Н. Усманова, Ф. А. Федоренко, “Безградиентные прокс-методы с неточным оракулом для негладких задач выпуклой стохастической оптимизации на симплексе”, *Автомат. и телемех.*, 2016, № 10, 57–77.

[4] O. Shamir, *An Optimal Algorithm for Bandit and Zero-Order Convex Optimization with Two-Point Feedback*, arXiv preprint arXiv:1507.08752.

[5] O. Shamir, “An Optimal Algorithm for Bandit and Zero-Order Convex Optimization with Two-Point Feedback”, *Journal of Machine Learning Research*, **18** (2017), 1–11.

[6] J. C. Duchi, M. I. Jordan, M. J. Wainwright, A. Wibisono, “Optimal Rates for Zero-Order Convex Optimization: The Power of Two Function Evaluations”, *IEEE Transaction on Information Theory*, **61**:5 (2015), 2788–2806.

[7] A. Blum, J. Hopcroft, R. Kannan, *Foundations of Data Science*, Vorabversion eines Lehrbuchs, 2016.

[8] K. Ball, “An elementary introduction to modern convex geometry”, *Flavors of Geometry*, **31** (1997).

[9] L. Bogolubsky, P. Dvurechensky, A. Gasnikov, G. Gusev, Yu. Nesterov, A. Raigorodskii, A. Tikhonov, M. Zhukovskii, “Learning Supervised PageRank with Gradient-Based and Gradient-Free Optimization Methods”, *NIPS 2016*, 2016.

[10] S. Boucheron, G. Lugosi, P. Massart, *Concentration inequalities: A nonasymptotic theory of independence*, Oxford university press, 2013.

[11] V. Milman, G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces. (With an Appendix by M. Gromov)*, Springer-Verlag, Berlin, 1986.

[12] В. А. Зорич, *Математический анализ в задачах естествознания*, МЦНМО, 2008.

[13] Баяндина А.С., Гасников А.В., Гулиев Ф.Ш., Лагуновская А.А., *Безградиентные двухточечные методы решения задач стохастической негладкой выпуклой*

оптимизации при наличии малых шумов не случайной природы, arXiv preprint arXiv:1701.03821.

[14] M. Lifshits, *Lectures on Gaussian Processes*, Springer, Heidelberg Dordrecht London New York, 2012.

[15] Э. А. Горбунов, Е. А. Воронцова, *Вычислительные эксперименты, иллюстрирующие явление концентрации равномерной меры на поверхности евклидовой сферы в малой окрестности экватора* https://github.com/evorontsova/Concentration-of-Measure/blob/master/Concentration%20of%20Measure.ipynb.

Э.А. Горбунов
Московский физико-технический институт, Москва
E-mail: ed-gorbunov@yandex.ru

Е.А. Воронцова
Дальневосточный федеральный университет, Владивосток
E-mail: vorontsovaea@gmail.com

А.В. Гасников
Московский физико-технический институт, Москва;
Кавказский математический центр, Адыгейский
государственный университет, Майкоп
E-mail: gasnikov@yandex.ru